Proceedings of the 52nd Hawaii International Conference on System Sciences | 2019

Narrowing the Theory's or Study's Scope May Increase Practical Relevance

Mikko Siponen
University of Jyvaskyla, Faculty of Information Technology P.O. Box 35, FI-40014 University of Jyvaskyla, Finland
mikko.t.siponen@jyu.fi

Tuula Klaavuniemi
Oncology Department, Southern Savonia Central Hospital
tuula.klaavuniemi@fimnet.fi

Abstract

Numerous articles in top IS journals note as a limitation and lack of generalizability that their findings are specific to a certain type of technology, culture, and so on. We argue that this generalizability concern is about limited scope (e.g., explanatory breadth). The IS literature notes this preference for generalizability as a characteristic of good science and it is sometimes confused with statistical generalizability We argue that such generalizability can be in conflict with explanation or prediction accuracy. An increase in scope (e.g., increasing explanatory breadth) can decrease explanation or prediction accuracy. Thus, in sciences such as cancer research, where explanation and prediction accuracy are highly valued, the cancer accounts (generally speaking) have become increasingly narrower (and less generalizable). IS thinking has not yet benefitted from these considerations. Whether generalizability is valued should be linked with the research aims. If the aim is practical applicability through explanation or prediction accuracy, then "limited" generalizability could be a strength rather than a weakness.

1. Introduction

The term generalizability is widely used in IS research articles, often in the limitations section, describing the lack of generalizability as a study limitation [1]. Typically, IS studies note that some of their findings are specific to a certain type of technology, country or organizational culture, and so on. For example, a study in India or Sweden may have India- or Sweden-specific findings, which are not applicable (generalizable) to outside of India or Sweden. Such India- or Sweden-specific findings are then noted as a limitation. As a concrete example, [2 p. 11] reported that "it is important to acknowledge several limitations. First, all data were collected from one organization, with one type of IT professional—the RW; hence, the results should not be generalized to other types of IT workers or organizations." To give another example, "Our findings should be interpreted in light of the limitations of this work. First, data were collected from two organizations that were of similar size and with similar operations in the same industry. Although this helped us control for possible industry differences, it limits the generalizability of our findings…Second, data were collected in the context of a specific ES implementation: a SAP ERP system. It is possible that our results would be different in other ERP systems.." [3 p. 1135].

Baskerville and Lee [1 p. 49] claimed that "It is incorrect and even harmful that many information systems researchers typically criticize their own intensive (qualitative, interpretive, critical, and case) research as lacking generalizability." Influential sources suggest that such lack of generalizability claims results because qualitative studies have to meet the standard of a statistical, sample-based conception of generalizability [4].[1] However, in the two examples above, the generalizability concern is not statistical generalizability; e.g., it does not relate to sample size adequateness. Findings regarding the "type of IT workers" are not generalizable to other IT workers [2]. In addition, "different ERPs" and "industry type" may not be transferable to other ERPs and industry types [3]. Nevertheless, in our experience, such a lack of generalizability can qualify as a reason to reject papers in top IS journals.

What do these examples demonstrate, if not issues of statistical generalizability? The point is not that theory contextualizations are not preferred. We suggest the interpretation that they stem from study, model, or theory scope preferences. For example, "an often admired quality of theories in natural sciences is their applicability to a range of settings" [5 p. 35]. Generalizability is also seen as the fundamental aim of

[1] "Published Research that Applies the Statistical, Sampling-Based Conception of Generalizability to Nonstatistical, Nonsampling Research" [4 p. 223]. Also: "researchers should not give up claims to generality on the basis of…small n," and they should have a "right to generalize and claim generality" [1 p. 61, 63].

URI: https://hdl.handle.net/10125/60060
ISBN: 978-0-9981331-2-6

sciences in IS [6]. This preference is also attributed to basic research and practical applicability [4 p. 221]: "the generalizability of an IS theory to different settings is important not only for purposes of basic research, but also for purposes of managing and solving problems that corporations and other organizations experience in society." Similarly, Weber [7 p. 15] noted that "Some theories cover a very narrow, constrained set of phenomena. Because of the limited range of phenomena it covers, however, it runs the risk it will be deemed uninteresting and unimportant." Moreover, Davison and Martinsons [8 p. 242] noted that "research that is relevant for and applicable to a wider range of phenomena is traditionally considered to be more useful." Finally, "Presumably, the wider the range of the theory's application, the more generalizability it offers and the stronger the theory" [9 p. 9]. We are not aware of any IS study that suggests that a narrower range makes a theory stronger or more useful.

These IS views make sense (with certain reservations) if the IS phenomenon is governed by laws. In the past, philosophers who assumed true laws, such as Fresner, Newton [10], and Lavoisier, outlined preferences according to which a wider explanatory breadth beat a narrower one. Moreover, Whewell, many figureheads of logical empiricism (e.g., Hempel), and analytical philosophers (e.g., Friedman and Kitcher) regarded the widening of explanatory breadth as an aim of science.

However, over the last 40 years or so, it has also become well documented in philosophy of science journals that generalizability issues are not that straightforward, even in the fields where we most expect them (e.g., in physics). Even the fundamental laws of physics might not be truly generalizable; they apply only in highly idealized counterfactual conditions [11]. Moreover, it is widely reported in many sciences that an increase in explanatory breadth (e.g., generalizability) often decreases explanation accuracy or prediction accuracy. Similarly, increasing realism tends to decrease explanatory breadth ("generalizability" decreased). For example, in cancer research, efforts to obtain increasingly better 1) explanation accuracy and 2) prediction accuracy have resulted in cancer accounts becoming increasingly narrower (progressively less generalizable). Moreover, highly specific treatments, which, in IS terms, would have a "lack of generalizability," have been introduced.

IS thinking has not yet benefitted from these "generalizability" considerations that have featured in many sciences. Several hundred papers have been written about this topic in philosophy of science journals under technical jargon, which may appear foreign to IS readers.

We explain some of these technical concepts to IS readers in anticipation that IS researchers will understand that narrowing scope is not tantamount to narrowing contributions. For example, if there is no need for explanation accuracy (or prediction accuracy) and a high level of realism, then the scope of the study or theory could be wide, and "lack of generalizability" can be seen as a demerit. However, if explanation or prediction accuracy is highly valued, then theories or studies may become increasingly narrower (and less generalizable) in indeterministic settings. In such cases, what has thus far been seen as "lack of generalizability" in IS could be seen not as a limitation, but as a merit of a study.

2. Preference for statistical generalizability or unification

Influential sources argue that qualitative studies are required to meet the standard of *statistical, sample-based conception of generalizability* [1, 4]. What could these sample-based generalization issues be? For example, you have 20 interviews, and the reviewers ask that you do (say) 100 more in order to have statistical power to claim (statistical generalizability) [12]. Lee and Baskerville [4] maintained that such a requirement for qualitative studies is unjustified. We agree that the requirement of statistical generalizability – e.g., high sample size – for all qualitative studies misunderstand the purpose of qualitative research. Having said that, we suggest that the evidence provided by Lee and Baskerville [4] points to concerns other than statistical generalizability (and sample size). Lee and Baskerville [4 p. 223] presented 12 examples of "Published Research that Applies the Statistical, Sampling-Based Conception of Generalizability to Nonstatistical, Nonsampling Research: Examples from Case Research." We do not doubt that this may happen. However, many (if not all) of Lee and Baskerville's [4 p. 223] examples could be interpreted as presenting cases in which the authors discussed *concerns other than* applying "the Statistical, Sampling-Based Conception of Generalizability to Nonstatistical, Nonsampling Research" [4]. For example, Robey and Sahay, in the example cited by Lee and Baskerville [4 p. 223], noted that "each context is different so we should expect different contextual elements… The findings should not even be extended to other settings... What is true for GIS in the two local country governments studied may be untrue for GIS in other governmental units or in private enterprise."

Robey and Sahay noted that "each context is different," and thus, "these findings should not even be extended to other settings." Robey and Sahay might not have referred to sample size adequateness; rather, "findings should not even be extended to other settings," and they may not refer to statistical generalizability[2]. Instead, "each context is different" implies that some *explanans*[3] are expected to be different per context. In the case of Robey and Sahay, organization-specific *explanans* clarify why findings "should not even be extended to other settings." The more organization-specific *explanans* there are, the fewer findings are transferrable or generalizable to all organizations. Discussion of organization (or context) specific *explanans* as a "lack of generalizability" is not necessarily the same thing as statistical generalizability. In IS, both concerns are discussed under generalizability. We suggest that in many of the examples by [4], the authors refer to *explanans*/*explanandum* differences and not necessarily "the Statistical, Sampling-Based Conception of Generalizability." The authors [4] and their critic [13] do not discuss this possibility.

A simple example hopefully clarifies the difference between these two types of generalizability claims[4]. Let us assume that we found that *explanans* A (e.g., a set of dynamic mechanisms) explains cancer subtype 1 (but not subtype 2), and *explanans* B explains cancer subtype 2 (but not subtype 1). In IS jargon, the findings from cancer subtype 1 are not generalizable to subtype 2 (and vice versa). However, this "lack of generalizability" has nothing do with the "Statistical, Sampling-Based Conception of Generalizability" [4]. Different cancers have different *explanans* (e.g., cancer subtype 1 is explained by different *explanans* from cancer subtype 2). If this is the case, then no statistical technique or increase in sample size can change this outcome. This is because the *explanans* are different, and an increase in sample size cannot make the *explanans* the same. Lee and Baskerville [36] seemed to discuss such issues as "Published Research that Applies the Statistical, Sampling-Based Conception of Generalizability to Nonstatistical, Nonsampling Research." In addition, numerous articles in MISQ applied qualitative and quantitative methods and discussed such context-specific findings as limitations and lack of generalizability. For example, findings on "type of IT workers" are not generalizable to other IT workers [2], and "different ERPs" and "industry type" may not be transferable to other ERPs and industry types [3]. In these examples, the generalizability concern is not sample size adequateness in claiming statistical power and generalizability. Why are such findings regarded as limitations and a "lack of generalizability" in IS rather than as strengths?

Assume that previous research had a cancer-type-specific theory (say, a lymphoma-specific theory) that explained all types of lymphomas; then, later, it was found that lymphomas could be classified as non-Hodgkin versus Hodgkin lymphomas [15, 16] that have different *explanans* and, therefore, subtype-specific treatments. In cancer research, this is understood as positive progress, not a limitation or "lack of generalizability." If it was found that patients with cancer subtype A are better off with specific treatment X, and cancer subtype B requires specific treatment Y, then describing such a finding as a lack of generalizability and a limitation would be odd, even though the results would be cancer-type-specific and could not be generalized to other cancers or even other lymphomas.

However, regarding such findings as implying a lack of generalizability or a limitation is understandable if there is *a priori* belief that a broad theory or study scope is better than a narrow one, as proposed by unificationists from Whewell [17] to Kitcher [18] (see next section). If the aim of the science is unifying (e.g., "finding common *explanans* across cancers, IT use or social media use") rather than dis-unifying (e.g., finding cancer-specific *explanans* or finding how the use of a certain type of computer game is different from a smartphone or a wide-screen computer), such views make sense. We have observed beliefs pertaining to scopes that are perceived as too narrow by top IS journal editors. For example, it was proposed by an SE that a study or theory that explained only Facebook (FB) use was not acceptable in a top IS journal because it did not explain other types of social media uses. Similarly, a study or theory of password selection was regarded as unacceptable because, for many reviewers/editors, "acceptable" studies or theories should explain a number of different IS security behaviors. These questions are often discussed as a lack of generalizability in IS. However, the concern is not about sample size or other statistical generalization issue, but rather whether FB use can be different from other types of social media use. FB use theory would have a narrower scope than a theory that explains all kinds of social media usages. Similarly, a

2 E.g., statistical generalizability "consists of generalizing from research findings about the sample to those same characteristics in the corresponding population" [13 p. 19]

3 Term *explanandum* (plural *explananda*) refers to a phenomenon that is explained [14]. *Explanans* (plural *explanantia*) explains the *explanandum* [14]. Originally, *explanans* were laws (and initial conditions) [14]. In contemporary philosophy of science, *explanans* can be other than laws, e.g., mechanism, factors, variables, causes, powers, events, process – virtually anything that accounts for the *explanandum*.

4 This is not an analogy between the Robey and Sahay example. The point is to illustrate how *explanans* differences are different from the "Statistical, Sampling-Based Conception of Generalizability".

theory/study on password selection would be narrower in scope than a theory/study that explains many kinds of IS security behaviors. These generalizability concerns implicitly assume unification preferences. We discuss these next.

3. Unification and laws

There are important differences between different sciences within the natural sciences and between the social and natural sciences. However, the IS discussions on (non-statistical) generalizability are not thus far related to such differences. The need for generalizability is justified or introduced as a general scientific preference. For example, IS views refer to "natural sciences" [5 p. 35] and "basic research" [4 p. 221]. Similarly, there is the claim that "generalizability should be given a higher position in the scientific process and the ultimate goal" is based on sciences in general rather than specific characteristics of IS or social sciences [6]. Finally, Davison and Martinsons [8 p. 242] noted that "research that is relevant for and applicable to a wider range of phenomena is traditionally considered to be more useful." Again, they [8] do not claim that this particular observation is specific to any particular science.

Since these IS views refer to natural science, basic research, and the goal of sciences in general, we shall now review what has been written on this topic in philosophy of science journals. The terms used are consilience of inductions, common cause, explanatory unification, explanation simplification, logical unification, derivational unification, value conflicts, etc. An overview of these concepts will shed new light on current IS practice and views on scope.

3.1. Laws and scope

It was once taken for granted that the scientific theories are (mainly) *true laws*. With the standard law concept, the issue of scope was straightforward: "Traditionally, the word 'laws' has been reserved for universally applicable, exceptionless generalizations" [19 p. 731]. For example, Popper [20] regarded theories as laws that were 100% exception-less [21; 19][5]. The paradigmatic case for laws was Newton's law of gravitation [22 p. 409]. Newton's theory "originally claimed to apply to all bodies through the universe at all times" [22 p. 409].

Many philosophers, who assumed that scientific theories are laws, outlined theory scope preferences, known as unification. But what is unification? This is what we discuss next. This discussion is important because many IS views on generalizability may implicitly or explicitly reflect this view.

Whewell [17] outlined a classical account "consilience of inductions." He claimed that "a test of the truth of the theory" is when a hypothesis explains more than "one class of facts" or when it predicts the "cases of a kind different from those which were contemplated in the formation of our hypothesis" [17]. For example, Whewell's [17] first requirement can be seen as unification and simplification [23]. Let us presume that different hypotheses explain Facebook (FB) use and Twitter use, and then, a new hypothesis (H3) explains both FB and Twitter use. Explanation simplification happens only if one hypothesis can explain both (instead of two). H3 has a wider scope than H1 and H2 in terms of the number of phenomena explained. Whewell's view can be interpreted in terms of simplification, common cause, or a hypothesis' scope or "explanatory breadth."[6]

Whewell's [17] doctrine was influential. A number of major thinkers, including Jevons, Fowler, Pierce, and Popper, agreed with Whewell that the best theory or hypothesis "is the one which has predicted new phenomena, explained phenomena of different kinds" [23 p. 177]. For example, Thagard [24 p. 30] reported "explanatory breadth" as a key criterion for evaluating theories, which means that "one theory has more explanatory breadth than its competitor if it explains more classes of facts." Numerous philosophers have outlined various unification theories, including Friedman's [25] explanatory unification, Mäki's [26] logical unification, Mäki's [27] derivational unification, and Kitcher's [18] explanatory unification. We can summarize these accounts with the following simplifying statement. **An increase in *explananda*** (number of phenomena explained or predicted) **increases the scope** (or breadth) **of a model or theory**. In this case, (non-statistical) generalizability increases. Let us assume that it is believed that a set of *explanans* A explains Twitter use and a set of *explanans* B explains FB use. Now, presume a new study that suggests that a set of *explanans* C explains both Twitter and FB use. With this new study, *explanandum* unification has occurred. Previously (in our example), Twitter and FB use was explained by a different set of *explanans*. This new study suggests the same

5 "Scientific law" means "Of all points in space and time (or in all regions of space and time) it is true that" [20 p. 68].

6 If H1 explains FB use and H2 explains Twitter use, then the new H3 seems to constitute both simpler explanations for the class "social media" but not for those of FB and Twitter. This is because both still have equally many numbers of explanatory reasons (hypotheses). However, in this simple example, H3 could be a common cause for FB and Twitter, while H1 and H2 could be separate causes. In this case, the "explanatory breadth" of H3 would be wider than those of H1 and H2.

explanans for Twitter and FB use. It can be regarded as introducing *explanandum* unification because it has unified two phenomena (Twitter and FB use), which were previously regarded as two different phenomena (because they had different sets of *explanans*: A and B). At the same time, the range of *explanandum* has increased with the new study. The new study can explain FB and Twitter use; while the previous accounts could only explain one.

In turn, **a decrease in explananda narrows the scope**, and (non-statistical) generalizability decreases. In the case of studies that explain, a decrease in *explananda* decreases explanatory breadth. Let us presume that according to paper 1, FB and Twitter have the same set of *explanans*. If one later shows in study 2 that they have different *explanans*, then dis-unification has happened. In addition, non-statistical generalizability has decreased from paper 1 to paper 2. That is to say that the scope of paper 2 is narrower than that of paper 1.

Classical unificationists from Whewell to Kitcher would not have appreciated what we have called a decrease in *explananda*. However, there are often other important concerns at stake, which were not taken into account in these classical unificationists' doctrines. Next, we explain what these concerns are. This discussion helps us understand that narrowing the scope (a decrease in *explananda* and a decrease in non-statistical generalizability) can be very important in many cases.

4. Value conflicts: accuracy versus scope

For Laudan [28], theories are solutions to important problems. Laudan [28 p. 35] introduced the terms comparative generality and weighting by generality: "if we can show for any two [scientific] problems p' and p, that any solution [i.e., theories] to p' must also constitute a solution for p (but not visa versa) then p' is more general, and thus of greater weight, than p." Laudan [28 p. 35] called this "comparative generality." His example is that finding the law for the motion of Mars has less "general comparatively" than finding the for the motion of all planets. However, he [28 p. 35] recognized that "there are many other cases which do not permit one to evaluate their comparative generality." What does Laudan mean by this? While he did not explain this, his former colleague provided an explanation that shocked many philosophers. We now turn to Kuhn.

Kuhn [29 p. 52] claimed that (what he called as) normal science, values "the steady extension of the scope." However, he maintained that on a rational basis, scientists even within one discipline could not agree on the precise meaning of this concept. Scientists might agree that a wide theory scope is better than a narrow one, but there was disagreement regarding what this meant precisely [29, 30]. Let us illustrate this idea with IS examples. Lee and Baskerville [4] discussed generalizability in terms of "different contexts." [31] discussed the importance of generalizability, which was termed "applicability to different environments" and "a variety of contexts." A Kuhnian might claim that scholars often agree with some concepts, say, that the study should be applicable to a "variety of contexts," but they cannot agree on a rational basis regarding when precisely a study meets these goals. For example, does a case study of two organizations meet a "variety of contexts" or "different contexts"?

Moreover, due to value conflicts, the issue becomes even more challenging, according to Kuhn [30 p. 262]:

"*In many concrete situations, different values, though all constitutive of good reason, dictate different conclusions, different choices. In such cases of value-conflict (e.g., one theory is simpler but one is more accurate) the relative weight placed on different values by different individuals can play a decisive role in choice. More important, though scientists share these values…, they do not all apply them in the same way. Simplicity, scope, fruitfulness, and even accuracy can be judged quite differently…by different people*".

Here, Kuhn introduced a term that has been very important in science, although we have not yet discussed it: accuracy (or precision). Before we discuss this concept, it is important to note that in biology, biochemistry, or cancer research, explanations are not primarily law-based, but mechanism-based [21, 32]. Mechanism-based accounts in biology and biochemistry are commonly regarded as "highly particularized" [21 p. 763]. Also, laws are questionable in social science [33].

Moreover, since the 1970s, philosophers have reported that even the fundamental laws of physics are not really true (exception-less) laws: "fundamental laws are not true, nor nearly true, nor true for the most part [11 p. 175]. For example, "Newton's first law…refers to what happens to a body that is subject to no external forces, but there are probably no such bodies" [34 p. 358]. As a final example, Nagel [35 p. 131] reported that "It is common if not normal for a theory to be formulated in terms of ideal concepts such as…perfect vacuum, infinitely slow expansion, perfect elasticity…" Ideal concepts are "simplifying false-hoods" [36 p. 242]. For example, assuming a perfect vacuum, which some theories of physics do, is a deliberate false representation of the phenomenon [37, 38]. As a result, even fundamental laws of physics either 1) make true claims, which apply only in highly idealized counterfactual settings, or 2) make false claims about how things are in actual settings [11]. For example, “how few are the known exact, true, and general laws that apply to actual as opposed to ideal conditions. They may be none at all” [19 p. 730]. Thus, “even our best theories [in physics] are severely limited in their scope.” [39 p. 13].

4.1 Science examples of accuracy versus scope

Explanation and prediction (accuracy) can be two different things. Explanation and prediction do not necessarily go hand in hand, a point not taken into account by, for example, the Hempel and Oppenheim [14] model. For instance, a prediction can be accurate, but an explanation might be less accurate or unknown. A case in point is that even though the detailed molecular mechanism of rituximab (a cancer drug) is not well understood, rituximab is a scientifically accepted treatment. The reason for this is related to the drug’s effectiveness. We might call that prediction accuracy. What does this mean? Take, for instance, an example of cancer treatment for one type of cancer: diffuse large B-cell lymphoma (DLBCL). One treatment combines rituximab and chemotherapy. For example, in DLBCL, at the 2-year follow-up point, 70% of patients treated with rituximab and chemotherapy were alive, compared to 57% who received only chemotherapy [40]. The prediction accuracy of chemotherapy in DLBCL can be said to be 57%, while that of rituximab combined with chemotherapy is 70%. However, prediction accuracy can be seen as a statistical average. In practice, this means that the effect of chemotherapy (treatments) varies from patient to patient. One might die before the two-year follow-up; another might be alive up to that point or even longer; for a third patient, the chemotherapy can have lethal side effects. Prediction accuracy rate studies may also vary from one study to another, for example, because cancers hardly follow laws [41], and cancer formation (carcinogenesis) is highly dynamic and random [42].

Kuhn’s reply to value conflicts shocked many philosophers [43]. Let us presume that two scholars agree that a “wide theory scope is important and (prediction) accuracy is important.” Let us presume that they do agree what both mean precisely, which Kuhn [29] denied. Let us say that the wide theory scope means a “theory that explains two types of cancers is better than a theory that explains one type of cancer.” Let us also presume that scholars agree on what accuracy means: it is the overall survival rate or the percentage of patients who are still alive after the (say) 2-year follow-up point (“prediction accuracy”). Let us presume that we have theory 1 (T1), which explains and has a treatment for two types of cancers A and B. Then we have two other separate theories (T2, T3) for cancers A and B. T2 can account only for cancer A (and nothing more), while T3 explains only cancer B (and nothing else). T1 is more generalizable; it explains cancers A and B, but its prediction accuracy is low. Say T1 has 20% accuracy for cancer A and 15% accuracy for cancer B. In turn, T2 and T3 are less generalizable. Each applies to only one type of cancer. But T2 has 90% accuracy for cancer A, and T3 has 90% accuracy for cancer B (90% of patients treated with T2- and T3-based treatments were alive at the standard two-year follow-up point).

Kuhn [30] claimed that in such circumstances, two scholars may diverge on which theory is better—even though they agree that a “wide theory scope is important and accuracy is important.” Even if they agree on what these two values mean precisely, which Kuhn once denied [29], they may still come to a different conclusion about which theory is preferred, according to Kuhn. For example, the first scholar prefers T1, while the other scholar prefers T2 and T3. For Kuhn, these are not a few exceptional cases. They “are what goes on in the sciences at times of theory choice” [44 p. 325]. Kuhn is highly skeptical that scholars can agree in such situations; the decision is personal and subjective. According to him, they cannot be solved by “evidence and reason” [43 p. 14]. We return to our cancer research example.

While T1 is more generalizable than T2 and T3, T2 and T3 have much better prediction accuracy than T1. Our readers may disagree with Kuhn—value conflicts could be solved. Most likely, oncologists would agree that T2 and T3 are better than T1, although they are less generalizable than T1. What happened here? R.M. Hare [45] would have called this an overriding of values. While breadth is an important value, in our example, prediction accuracy is even more important. In our example, we cannot have both, and prediction accuracy overrides generalizability (explanatory breadth). Cancer researchers and oncologists would not necessarily trade off a specific theory to have a general theory of cancer if the general theory is less accurate. However, we argue that resolving such value conflicts may be situational and dependent on the aim of the research. In mathematical models of cancer research, generalizability could be more important than an accurate (and realistic) explanation or prediction, especially when there is no direct expectation to apply the research in practice.

Similar examples of the resolution of value conflicts can be given in many sciences, but Kuhn [29, 30, 44] was not aware of this. In population biology, MacArthur and Levins [46] traded off precision in order to have better realism and wider generality. Others [47, 48] traded off realism to have wider generality and greater precision by including knowingly unrealistic assumptions in their models. This was also the approach in Nobel laureate Friedman's [49] methodology of economics. He [49] felt that he could not have a wide scope, realism, and accurate predictions at the same time. For him [49], a wide scope and prediction accuracy were much more important than realism, which he traded off. Thus, for him, good theories in economics can have assumptions that "never are" realistic. He purposefully traded off realism for a wide scope and prediction accuracy. His exemplar was physics, in general, and Galileo's law of fall in particular [38]. Friedman [49] knew that the law has numerous purposefully false assumptions.

Had he looked up cancer research as an exemplar instead of Galileo's law, he might not have traded off realism so easily because it would have resulted in cancer treatments that do not work in practice. Recall that mechanism-based explanations are commonly regarded as highly particularized in biology [21]. As an example of a cancer mechanism in biology and biochemistry, the humoral theory once explained all medical concerns from cancer to melancholy [50]. Thereafter came cancer-specific accounts, which explained not all medical diseases, but all cancers. Later, upon realizing that different cancers had different complex dynamic mechanisms, cancer research followed up, and narrower cancer-type-specific explanations were sought [50]. However, cancer research did not stop there. There are now hundreds of distinct cancer accounts for different types of cancers [50]. There are cancer-specific explanations, which may not be generalizable to any other cancers or phenomena in science, according to current knowledge. Similarly, there are cancer-specific drugs that are applicable only to certain types of cancers, and their effect may not be equally generalizable to other types of cancers [50]. Such particularistic theorizing in cancer research has, on its own, saved the lives of numerous people. If scientific studies or theories must be applicable to various different settings (e.g., explain many types of cancers) or always toward increasingly greater scope in order to prove that the study is good science, then cancer-type-specific theorizing exemplifies not-so-good science. The progress of cancer research in biology and biochemistry can be (partly) explained by prioritizing explanation accuracy over generalizability.

5. Discussion

A decrease in or narrowness of theory scope does not necessarily reduce practical and scientific relevance. Many influential IS scholars, regardless of background, are certain that a wider scope is better than a narrow scope, both scientifically and practically. Alternatively, they suggest that a narrow scope is *less useful*, *less strong*, *more uninteresting*, and *more unimportant* than a wider scope [8 p. 24] [9 p. 9] [7 p. 15]. Such views make sense in the logical empiricist and Popperian utopia where there are true, universal (exception-less) laws. The extent to which such laws exist in physics is questionable, let alone in medical research, biochemistry, or social sciences.

Outside of such a utopian world, narrowing the theory or study scope does not necessarily decrease practical importance or impact. This becomes clear when we discuss how potential scientific impacts can be different from theory scope generality and specificity as well as how common a problem the issue (*explanandum*) is. For example, pancreatic cancer is rarer than breast cancer. However, currently, pancreatic cancer is a more serious problem, on average, than breast cancer because the former is lethal in nearly all cases (breast cancer is more often curable, if found early enough). We may say that the impact of pancreatic cancer is higher than that of breast cancer, if the yardstick is lethality. In addition, the scope of generality and specificity does not necessarily go hand in hand with the scientific impact. Arguably, there is something unique in the *explanans* of pancreatic cancer that is not fully understood because it cannot be cured. Who wants to claim that we should not examine lethal pancreatic cancer even though its scope is narrow, even though it has unique *explanans* and yet is rare? It is also possible that the theory with a highly narrow scope can have greater potential for practical impact than a study with a more general scope. Many scientists have recognized that narrowing the theoretical scope may increase practical relevance. In IS jargon, narrowing (explanatory) the scope is not tantamount to weakening contributions. What, therefore, indicates a narrowing scope? We discuss two important reasons next.

Increased explanatory accuracy is one reason why a scope is narrowed. We link explanatory accuracy with *explanans*. For example, non-Hodgkin and Hodgkin are two types of lymphomas, but why do they require separate theorizing? The answer is because these two lymphomas have different *explanans*. If two *explananda* can be adequately understood using the same *explanans*, then they are most likely about the same phenomenon. If two *explananda* have different *explanans*, then this is an indication that these are two separate phenomena, which may require distinct theorizing. We can have a theory that explains both non-Hodgkin and Hodgkin lymphomas, but this theory would be less accurate than specific theories for each. Moreover, we also need studies with a narrower scope, even when such differences in *explanans* may not exist. For the sake of simplicity, let us presume that all lymphomas are either accounted by Theory (T) 1 or T2. This we could only know by scrutinizing all candidates for lymphomas. This idea (differences in *explanans*) also applies in IS. If people have different reasons (*explanans*) for using (say) Facebook rather than (say) Twitter, then this is a good indication that separate theorizing may be required for each. In such a case, we can have a more abstract and general theory that explains both Facebook and Twitter use, but this theory may be less accurate than specific theories for each. Moreover, the problem with the general theory may be more than an accuracy issue. For example, models of computer abuse explain all types of information security policy non-compliance issues, from not locking a computer to reusing the same password across different accounts. However, let us presume that reusing a password is explained by problems with memorizing, while other types of information security policy non-compliance issues are not. If this were the case, then models of computer abuse would not only be inaccurate, but mainly inappropriate to explain password reuse behavior.

Increased prediction accuracy is another reason why a scope is narrowed. In medical research, prediction accuracy is linked with the effect of the treatment. Lymphoma-type specific or unique treatments have a narrower scope than a "general treatment" for all lymphomas. Why are such specific treatments developed and used? The answer is straightforward. It is hoped that they will be more effective than the general treatment. In fact, a general treatment for lymphomas can even be life threatening for treating specific lymphomas.

What does this means for IS? We could say that "the more effective the treatment, the better," and this a posteriori matter then ultimately sets the scope. However, the issue is far from being that simple. For at least two complex reasons, there is hardly a single concrete solution that is universal in every IS research area. First, what counts as "effectiveness" (prediction accuracy) varies even within medicine. Effectiveness means different things in cancer treatments compared with (say) anesthesiology treatments. Similar differences are assumed in different areas of IS. For example, a key area in IS security is improving users' behavior through interventions such as fear appeals. But what counts as effectiveness is not straightforward. For instance, are we measuring the effect right after the intervention, two weeks later, or two months later? Do we count the possible side effects of fear appeals (e.g., users stop reading security-related email)? Moreover, not all studies can do all of these things, and such decisions implicitly limit the scope. Finally, the same measures of effectiveness may not work with (say) IT use. These issues must be discussed carefully in future IS research.

Second, prediction accuracy is also a complex methodological issue. For example, prediction and explanation accuracy of the same phenomenon are typically importantly different in actual settings versus 1) counterfactual laboratory settings, 2) mathematical models, or 3) statistical models with statistical averages. A case in point are fundamental laws of physics. They commonly contain purposeful false assumptions (e.g., perfect vacuum, magnetic, and other forces are absent; air pressure is nil) [11, 38]. Their predictions may hold in counterfactual settings (e.g., in laboratory, mathematical models), where these assumptions could be met. However, in real settings (generally speaking), their predictions can be, strictly speaking, false [11, 51][7]. In other words, the scope of a law in physics can be made purposefully wide in a counterfactual model (e.g., making all forces absent), with full awareness that the model falsely describes the actual phenomenon. Thus, the scope of a theory in a model or theoretical settings can be wide. In real settings, the scope of the same theory can be either highly limited or even inapplicable (without serious modifications). Such issues also exist in IS, albeit they might not have been recognized. These issues have scope implications and must be discussed and debated in future IS research.

6. Summary

Influential IS scholars outline an a priori belief: wider scope beats narrower scope, or studies with narrower scope are less strong, less useful, less interesting and less important than studies with wider scope. If we evaluate cancer research in light of these IS views, then both cancer biology and medical oncology have become (generally speaking) increasingly less useful, less strong, more uninteresting, and more unimportant over the last 200 years. Such beliefs related to IS scope should be rejected in their original form. However, they are understandable, if one's understanding of the philosophy of science is based on the arguments of logical empiricists (e.g., Hempel) or Popper up to the 1970s. In such a worldview, scientific theories were true exceptionless laws, and singular events or observations were explained by these laws. Today, it is known that even the fundamental laws of physics do not constitute such laws [11]. In cancer research, for example, a phenomenon can be rare and have a narrow

[7] "The models that our theories [of physics] are able to handle are deliberate falsifications of reality…the theory may then accurately describe the workings of the model, but the model does not describe the phenomena." [51 p. 200]

scope, yet a study of it can be scientifically and practically relevant and valuable. Narrowing the scope does not necessarily decrease the contribution (scientific or practical impact); it may increase it. However, we are not saying that a narrower scope is *a priori* preferred over a wider one. Our view is that in empirical sciences (outside of logic/mathematics), theory/study scope is not an *a priori* issue, but largely an *a posteriori* issue. Having said that, a theory or study scope may also be narrowed or widened in sciences for methodological or instrumental reasons. For example, increasing the scope (generalizability) in sciences (from physics to economy) is often done at the expense of realism and accuracy by 1) omitting relevant *explanans* (method of isolation), 2) having a less accurate description of the *explanans* (method of abstraction), or 3) introducing purposeful falsehoods in the *explanans* or *explanandum* (idealizations). These have scope implications, and future IS philosophy is needed to understand all of these methods, including prediction accuracy, and how they are implicitly or explicitly used in IS.

7. References

[1] Baskerville, R. & Lee, A.S. (1999) Distinctions among different types of generalizing in information systems research. In O. Ngwenyama et al. (eds.), New Information Technologies in Organizational Processes. IFIP International Federation for Information Processing.

[2] Ahuja, M, Chudoba, K, Kacmar, C., Mcknight, H. & George, J. (2007) IT Road Warriors: Balancing Work-Family Conflict, Job Autonomy, and Work Overload to Mitigate Turnover Intentions. MIS Quarterly, 3, 1-17.

[3] Bala, H. & Venkatesh, V. (2013) Changes in Employees' Job Characteristics During an Enterprise System Implementation: A Latent Growth Modeling Perspective. MIS Quarterly 37, 1113-1140.

[4] Lee, A. & Baskerville, R. (2003) Generalizing Generalizability in Information Systems Research. Information Systems Research 14, 221-243.

[5] Lee, A. (1989) A Scientific Methodology for MIS Case Studies. MIS Quarterly 13, 33-50.

[6] Cheng A., Dimoka, A and Pavlou, P (2016) Context May Be King, but Generalizability Is the Emperor! Journal of Information Technology, 31, 257–264.

[7] Weber R, (2012), Evaluating and Developing Theories in the Information Systems Discipline, Journal of the Association for Information Systems, 13 (1): 1-30.

[8] Davison, R.M. & Martinsons, M.G. (2016) Context is King! Considering Particularism in Research Design and Reporting. Journal of Information Technology, 31, 241-249.

[9] Hassan, N. and Lowry, P. 2015. "Seeking Middle-range Theories in Information Systems Research," International Conference on Information Systems 2015, December 13–18.

[10] Halonen, I. & Hintikka, J. (1999) Unification – It's Magnificent but Is It Explanation? Synthese 120, 27–47.

[11] Cartwright, N. (1983) How the Laws of Physics Lie. Oxford: Clarendon Press.

[12] Aguirre-Urreta, M & Rönkkö, M. (2015) Sample Size Determination and Statistical Power Analysis in PLS Using R: An Annotated Tutorial. Communications of the Association for Information Systems 36(3).

[13] Williams, J. N., & Tsang, E. W. (2015). Classifying generalization: paradigm war or abuse of terminology?. Journal of Information Technology, 30(1), 18-29.

[14] Hempel C.G. & Oppenheim, P. (1948) Studies in the Logic of Explanation. Philosophy of Science, 15, 135-175.

[15] Bödör C. & Reiniger L. (2016) Catalog of genetic progression of human cancers: non-Hodgkin lymphoma. Cancer Metastasis Rev. 35, 109-27.

[16] Küppers, R., Engert, A. & Hansmann, ML (2012) Hodgkin Lymphoma. The Journal of Clinical Investigation 122, 3439-3447.

[17] Whewell, W. 1840. Philosophy of the Inductive Science.

[18] Kitcher, P. (1981) Explanatory Unification. Philosophy of Science 48, 507-531.

[19] Teller, P. (2004) The Law-Idealization. Philosophy of Science 7, 730-741.

[20] Popper K. R. (1968) The Logic of Scientific Discovery, New York: Harper Torchbooks.

[21] Bechtel, W. (2009) Generalization and Discovery by Assuming Conserved Mechanisms: Cross-Species Research on Circadian Oscillators. Philosophy of Science 76, 762–773.

[22] Darden, L. (1996). Generalizations in Biology. Stud. Hist. Phil. Sci 27(3):409-419.

[23] Laudan, L. (1981) William Whewell on the consilience of inductions. In L. Laudan (eds): Science and hypothesis, pp. 163-180.

[24] Thagard, P. 2010. "Evolution, Creation, and the Philosophy of Science". In Epistemology and Science Education: Understanding the Evolution vs. Intelligent Design Controversy, ed. Roger S. Taylor and Michel Ferrari, 20-37. Milton Park: Routledge.

[25] Friedman, M. (1974) Explanation and Scientific Understanding, Journal of Philosophy 71, 5-19.

[26] Mäki, U. (1990) Scientific realism and Austrian explanation. Review of Political Economy 2, 310-344.

[27] Mäki, U. (2001) Explanatory unification: double and doubtful. Philosophy of the Social Sciences, 31, 488-506.

[28] Laudan, L. (1978) Progress and its problems: Towards a theory of scientific growth. University of California Press.

[29] Kuhn, TS (1962) The Structure of Scientific Revolutions. University of Chicago Press.

[30] Kuhn, T.S. (1970). Reflections on my critics. In I Lakatos and A Musgrave (eds), criticism and the growth of knowledge. Cambridge University Press.

[31] Hevner, A., March, S., Park, J. & Ram, S. (2004) " Design Science in Information Systems Research. MIS Quarterly, 28, 75-105.

[32] Thagard P. (2003) Pathways to Biomedical Discovery. Philosophy of Science 70, 235-254.

[33] Hedström, P & Ylikoski, P. (2010) Causal Mechanisms in the Social Sciences. Annual Review of Sociology 36:49-67.

[34] Ladyman, J. 2008. Idealization. In Martin Curd, Stathis Psillos (eds) The Routledge Companion to Philosophy of Science.

[35] Nagel, E. (1961). The Structure of Science. Hackett publishing company.

[36] Levy & Bechtel 2013. Abstraction and the Organization of Mechanisms, philosophy of science 80: 241–261.

[37] Giere RN, (1988) Explaining Science: A cognitive approach, The University of Chicago Press.

[38] Mäki, U. 2009. Realistic realism about unrealistic models. In Kincaid, H., & Ross, D. (Eds.), The Oxford Handbook of Philosophy of Economics: 68-98. New York: Oxford University Press.

[39] Cartwright, N. (1989) Nature's Capacities and their Measurement. Oxford: Clarendon.

[40] Coiffier, B., Lepage, E., Brière, J., Herbrecht, R., Tilly, H., Bouabdallah, R., Morel, P., Van Den Neste, E., Salles, G., Gaulard, P., Reyes, F., Lederlin, P., Gisselbrecht, C. (2002). CHOP Chemotherapy plus Rituximab Compared with CHOP Alone in Elderly Patients with Diffuse Large-B-Cell Lymphoma. New England Journal of Medicine, 346 (4): 235–242.

[41] Thagard P. (1998) Explaining disease: Correlations, Causes, and Mechanisms. Minds and Machines 8, 61-78.

[42] Tomasetti C. & Vogelstein B. (2015) Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science. 2015 2;347, 78–81.

[43] Laudan, L. (1996) Beyond Positivism and Relativism: Theory, method, and evidence. Boulder: Westview Press.

[44] Kuhn TS 1977. The Essential Tension: Selected Studies in Scientific Tradition and Change. University of Chicago Press, US.

[45] Hare, R.M. (1963) Freedom and Reason. Oxford University press.

[46] Macarthur, R & Levins, R. (1964) Competition, habitat selection, and character displacement in a patchy environment. Proceedings of the National Academy of Sciences 51(6): 1207-1210.

[47] Holling, C. (1959) The Components of Predation as Revealed by a Study of Small-Mammal Predation of the European Pine Sawfly. Canadian Entomologists 91(5): 293-320.

[48] Watt, K. (1956) The Choice and Solution of Mathematical Models for Predicting and Maximizing the Yield of a Fishery. Journal of the Fisheries Research Board of Canada 13, 613-645.

[49] Friedman, M. (1953) The Methodology for Positive Economics. In: Essays in Positive Economics, pp. 3–34, Chicago.

[50] Sudhakara, A. (2009) History of Cancer, Ancient and Modern Treatment Methods. Journal of Cancer Science & Therapy 1(2): 1–4.

[51] Chalmers A (1993) So the laws of physics needn't lie, Australasian Journal of Philosophy, 71, 196-205.